DOCUMENT RESUME

ED 217 870 IR 010 256

AUTHOR Reeves, Thomas C.; Lent, Richard M.
TITLE Levels of Evaluation for Computer-Based Instruction.
PUB DATE 20 Mar 82
NOTE 29p.; Paper presented at the Annual Meeting of the American Educational Research Association (New York, NY, March 20, 1982).

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Computer Assisted Instruction; Costs; *Documentation; Evaluation Criteria; *Evaluation Methods; *Formative Evaluation; Observation; *Program Development; Recordkeeping; Self Evaluation (Groups)
IDENTIFIERS *Immediate Learner Effectiveness; *Impact Evaluation; Operational Testing

ABSTRACT

The uses and methods of four levels of evaluation which can be conducted during the development and implementation phases of computer-based instruction (CBI) programs are discussed in this paper. The levels of evaluation presented are: (1) documentation, (2) formative evaluation, (3) assessment of immediate learner effectiveness, and (4) impact evaluation. The components of the documentation level of evaluation, which include monitoring project costs, recordkeeping, and personal observation, are outlined; the processes involved in formative evaluation, including internal review and operational testing, are described; the various factors involved in the assessment of immediate learner effectiveness are summarized; and the nature of impact evaluation is discussed. A set of nine sample evaluation forms relating to different levels of evaluation and their associated methods and a 24-item bibliography accompany the text. (JL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED217870

Levels of Evaluation for Computer-Based Instruction

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.

Minor changes have been made to improve reproduction quality.

Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

Thomas C. Reeves, Ph.D.
Richard M. Lent, Ph.D.

Center for Instructional Development and Evaluation
University of Maryland University College

A Paper Presented at the Annual Meeting of the
American Educational Research Association
New York, March 20, 1982, Session 15.04

For further information:
Thomas C. Reeves
Richard M. Lent
Center for Instructional Development
and Evaluation
University of Maryland University College
University Boulevard at Adelphi Road
College Park, Maryland 20742

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Thomas C. Reeves
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

IR010252

LEVELS OF EVALUATION FOR COMPUTER-BASED INSTRUCTION

by
Thomas C. Reeves, Ph.D.
Richard M. Lent, Ph.D.

Center for Instructional Development and Evaluation
University of Maryland University College

The purpose of this paper is to present the uses and methods of four levels of evaluation which can be conducted during the development and implementation of computer-based instruction (CBI). The overall goal of evaluation during the development and implementation of CBI (or any other program) is to provide decision-makers such as sponsors, administrators, developers, programmers, and instructors with timely, accurate information which will contribute to decisions about the improvement, continuance, and/or expansion of the program (Anderson & Ball, 1978). To accomplish the goal of providing a CBI project's sponsors and personnel with information for making decisions, four levels or types of evaluation are recommended: documentation, formative evaluation, assessment of immediate learner effectiveness, and impact evaluation.

Before describing the uses and methods of these four levels of evaluation, it should be noted that a review of the CBI literature reveals that very few CBI projects have employed more than one or two of these levels. For example, Fitzpatrick and Howard (1976), Muston and Wagstaff (1976), and Rubin, Geller, and Hanks (1977) all concentrated their evaluation efforts at the formative level, using the ubiquitous student questionnaire. Fletcher and Suppes (1976), Su and Eman (1975), and Swigger (1976), on the other hand, spent their evaluation dollar in attempts to assess the immediate learner effectiveness of their CBI, utilizing tests and quasi-experimental designs. The authors of this paper will not argue that every CBI project should include all four levels of evaluation, but do maintain that CBI projects would benefit from a wider consideration of the decision-related information needs of their participants and consequent fuller representation of the evaluation levels described below.

Level 1 - Documentation

In the simplest sense, documentation involves keeping records of when and where various project activities occur, what the activities cost, and who participates in them. Documentation is important for at least two reasons: administrators must have documentation data to account for the use of project resources and documentation data is required for the other three levels of evaluation described in this paper.

It is easy for participants in a CBI project to become so involved in the creative and logistical aspects of the project that basic documentation data about the design and development of CBI programs, and who uses them, when, where, for how long, at what rate, and so forth can be uncollected or lost. To avoid these problems, computer-based data collection and data management systems and simple record keeping instruments are recommended.

Documentation of a CBI project's activities is a fundamental part of the evaluation effort regardless of the particular stage that the project is in: from design and development, to production and delivery, and finally to implementation and long-term operation of the instructional system. Economists refer to these stages as a project's "life cycle." Is important for the evaluator to document aspects of each of these stages because each involves its own subset of inputs, processes and outcomes which ultimately will add up to the project's overall results. Furthermore, it is only by tracking the pattern of decisions, events, resources and activities (of both project staff and student users) throughout the project's life that explanations for the nature of the final results can be developed. Some specific aspects of documentation - level evaluation activities appropriate to all states of a project's life are described here.

Documentation of Project Costs

During all stages of the project it is important to gather data on the type, cost and rationale for the resources being consumed by the project. A CBI project represents a special investment of resources to develop an approach to an instructional problem. On a short term basis, the amount of resources expended on a CBI project is not likely to be justified by the nature of the immediate benefits received. Instead, the hope is that the initial research and development investment which a CBI project represents will ultimately be justified by the long-term benefits to be received as the project's products (materials and ideas) are put into use over time. For example, funds spent to develop CBI which teaches the handicapped about career opportunities may seem extravagant in the age of "Reaganomics," but completely justified if an unintended outcome of the project is the discovery of new technologies through which handicapped persons may interact with computers.

It is beyond the intent and scope of this paper to present a thorough delineation of the methods of conducting cost-effectiveness analyses of instructional technology projects. The interested reader is referred to Beilby (1979), Felden and Pearson (1978), and Lent, (1979, 1980). However, it is important to point out that such analyses are impossible if project costs are not documented throughout the life of the project. Figure 1 presents a general model for the analysis of project resources across a project's life cycles. Each cell of the model indicates a particular type of resource likely to be consumed by the

project. The amount and value ($) of each resource being consumed must be documented and its properties as a fixed or variable, and recurring or nonrecurring cost of the project determined. Figure 2 presents one of the fundamental data collection instruments used to obtain the detailed information necessary for useful cost analyses of instructional technology projects. The "activity log" shown is a form used by each member of the project to log his/her time and to assign it to a major cost center and a specific development, production or operation stage of project activity. Such data is periodically entered into a computer-based management system and will ultimately be used to determine the extent and allocation of personnel resources across the matrix in Figure 1.

Automatic Record Keeping

One of the major strengths of CBI is that the same computers which "deliver" the instruction can also be programmed to collect most of the evaluation data required for project documentation. However, the unobtrusive collection of documentation data must be carefully planned and meticulously incorporated into the instructional programs. To begin, a complete list of the types of documentation data to be collected by the computer must be constructed. This list will be derived from the decision-based information needs of the project participants. Examples of such data include demographic data (names, ID numbers, educational status), time-on-system (sign-in/sign-off), student responses (tests, questionnaires) and curricular progress status. Time-on-system is an especially important type of data because it is required for the computation of other data, e.g., student rate of lesson completion, and because time-on-task has been demonstrated to be a powerful explanatory variable in research and evaluation of instructional systems (Rosenshine, 1977). Therefore, a real time clock mechanism will be a necessary component of the system hardware of most CBI projects.

The inclusion of record keeping routines in CBI will enable project personnel to obtain accurate answers to decision-oriented questions at any time. For example, timely information about system utilization at different sites should permit project administrators to make better decisions about project site changes. Similarly, rapid turn-around of student response data may allow program developers to make critical adjustments in the instructional programs.

If microcomputers are used in a CBI project, it may be necessary to arrange for the up-loading of the micro-collected data to a larger computer for analysis using standardized data analysis programs such as SPSS (Nie, Hull, Jenkins, Steinbrenner, & Bent, 1975). Such a procedure will also lessen the drain on the micro's memory caused by the data collection and maintenance programs.

It is critical that computer-based data collection be as unobtrusive as possible. This is especially important when the intended users are children or adult learners who may be easily frustrated by tedious questions and responses. The simulation power of computers, especially when integrated with videodiscs, can be used to collect data in an interesting and even motivating manner. For example, a CBI project presently under development at the University of Maryland uses an Adventure-like simulation to obtain necessary demographic data. Students enter an unusual building and go from floor to floor in a simulated elevator to sign up as members of an elite time-traveling force. During this brief but stimulating simulation, students not only provide the computer with the required demographic data, but they also learn to use the student response components of the microcomputer system, viz., the keyboard and touch-sensitive screen.

Personal Observation

Not all the data for complete and meaningful documentation of a CBI project can be collected by computer-based programs. Human beings must collect and report data concerning external factors which may affect the implementation of CBI such as power failures, equipment breakdowns, and fluctuations in the available student population. Figure 3 is a Site Manager`s Incident Report form used to record incidents which may have powerful explanatory power in evaluating the outcomes of a particular CBI project. Another important type of documentation data is the eyewitness account of the processes and outcomes of the project. Figure 4 is an Anecdotal Record form to be used by project participants to record the incidents observed by or reported to them. These "stories" will often make an otherwise cold and impersonal evaluation report an alive and human source of decision-based information.

To summarize, the first level of evaluation, documentation, involves the systematic recording of all the "who", "what", "when", "where", and "for how much" data required for accounting for the use of project resources, making information-based decisions, and/or carrying out higher levels of evaluation. In the process of documentation, the following recommendations should be heeded:

1. Cost analysis is a fundamental part of any project evaluation. A useful cost analysis requires a thorough documentation of all resources consumed by the project through all stages of the project's life. Furhermore, these resources must be further identified as to whether they represent fixed or variable and recurring or nonrecurring expenditures for a project of this type.

2. Whenever possible, the computer itself should be used to collect documentation data, especially if it can be done unobtrusively.

3. Human data collection, employing simple forms, is necessary whenever an evaluation is intended to inform its audiences about the personal or qualitative "story" of the development, implementation, or outcomes of a CBI project.

Level 2 - Formative Evaluation

Formative evaluation, which is the collection of the opinions, suggestions, and criticisms of project participants, is of critical importance to the revision and improvement of a CBI project. As with the first level of evaluation, the types of formative evaluation activities will differ according to whether the project is in the development or implementation phase. Formative evaluation methods such as expert review and pilot tests are integral parts of the lesson design and development process followed in any CBI project. And yet, as critical as formative evaluation is during the development phase, it cannot be de-emphasized during the implementation phase. Even the most rigorous design and development process cannot yield CBI programs immune to installation and implementation problems.

Sometimes formative evaluation is belittled as being not much more than a "smiles test" to find out whether or not program users like an education or training activity. However, formative evaluation methods such as questionnaires. interviews, observation, and expert review can be used to collect data much more useful than simply the degree of program acceptance. Information about the appropriateness of program objectives, content, instructional methods, sequence, and pace can be collected by formative evaluation methods and used to increase the effectiveness and efficiency of CBI.

Two major approaches to formative evaluation are internal review and operational testing. Internal review is the process of systematically reviewing the content and instructional processes of CBI before it is put into operation. Internal review includes such activities as small scale pilot tests and expert review. Operational testing refers to the process of collecting information for improving instruction during and after its actual implementation. Operational testing employs such methods as computer-based questionnaires, personal interviews, and observation.

Internal Review

It is true that most CBI is developed using some sort of systematic approach involving needs assessment, specification of goals and objectives, instructional design and production, and program evaluation. Models for developing CBI include many suggestions for formative evaluation. (See, for example, O'Neil, 1979 or Rahmlow, Fratini, Ghesquiere, 1980.)

Internal review can be relatively informal or highly systematic. An informal review might consist of development staff or content experts addressing questions such as:

a. Is the instructional content accurate? Content experts can be hired to periodically review instructional content, both before and after it is programmed.

b. Are the principles and practices of effective CBI incorporated into lessons? Figure 5 is a CAI Readiness Checklist developed by McPherson-Turner (1979) for review of CBI. Such a checklist should be regularly used by instructional developers to review the quality of their products.

c. Does the instruction account for all the planned outcomes? Instructional experts can often detect whether all the intended outcomes for a particular sequence of CBI are actually addressed during the activity. This may at first seem self-evident, but it is not uncommon for developers of CBI to have objectives (particularly affective ones) for which no instructional activities are designed. For example, CBI can be designed to teach new accounting procedures adequately, but may fail to motivate the learners to implement the changes.

d. Is the activity interesting? All too often, CBI takes the form of monotonous sequences of pretests, lessons, and posttests. If the power of the computer to simulate is used adequately, learners can be put into situations in which they want to learn. For example, if you wish to teach functionally illiterate soldiers how to use a military indexing system, you can simulate a unusual and entertaining "manual" about sports, cars, or other interests which soldiers will want to read. They will learn the appropriate reference skills in order to progress further in the simulation.

A more systematic review of CBI would involve a small scale pilot test with a sample of potential users. Learner performance and reactions as well as observers' evaluations can provide information concerning necessary revisions of the inputs, processes, and effects of CBI. The feedback from users and observers can provide information needed to address questions such as:

a. What knowledge, skills, and attitudes are prerequisite to the CBI?

b. What parts of the CBI work and which don't?

c. What intended outcomes are accomplished?

d. What unintended outcomes are detected?

A primary source for formative information during a pilot test is participant post-instructional performance on measures of the intended outcomes. However, this is only true when the performance measures or tests are carefully constructed and well validated. CBI rarely enjoys the benefits of reliable and valid outcome measures, and other data sources such as student reactions and observer reports must be the basis for program revision.

If, on the other hand, reliable and valid tests of student performance are available, they can provide excellent information, especially if the links between program objectives, instructional processes, and outcome measures have been carefully explicated. Once again, the power of computers to simulate reality can frequently be used. For example, CBI can be developed to teach CPR using a programmed manikin. If participants in pilot tests of this instruction have difficulty with a particular step as indicated by performance measures involving the dummy, the CBI designed to teach that step can be augmented or revised.

A second source for formative evaluation data during a pilot test is user reaction to the program. During operational testing student reaction is usually collected by computer-based questionnaires, but an individual or group interview is a better method in a small scale pilot test. These interviews should be open-ended and participants should be encouraged to be as critical as possible. Participants in pilot tests can also be asked to record their reactions as they progress through CBI, noting the high and low points and making suggestions for improvement.

A third valuable source of information during a pilot test can be observers who are experts in either the content of the lessons or the processes of CBI. Even a non-expert observer can be useful, especially in interacting with the learner to point out "leaps" and "gaps" in the lessons or unclear content presentations. As with all the procedures during a pilot test, emphasis during observations should be placed on diagnosing problems and suggesting remedies.

In summary, before an educational activity is actually implemented it should be subjected to internal review procedures including, if possible, small scale pilot tests with potential users. Unfortunately, few CBI development projects have the budget or timeline to spend large amounts of money and time on these procedures. Realistically, much formative evaluation will be conducted during or after the CBI is actually implemented using the techniques of operational testing.

Operational Testing

Operational testing is the process of collecting data for the improvement of CBI during or after its actual implementation.

The major sources of information during operational testing are a. measures of participant performance, b. computer-based questions and personal interviews about user reactions, and c. observations by designers and other observers. Operational testing is essential because CBI is usually developed to be used in many different settings over a relatively long period of time. As noted earlier, even the more rigorously designed and developed CBI can be improved after it is installed.

In operational testing, user performance is an essential variable for scrutiny if and only if reliable and valid measures of performance are used. More will be explained about the issues of the reliability and validity of computer-based tests in the next section of this paper concerning "Assessment of Immediate Learner Effectiveness." What is important to understand now is that, given good measures of the outcomes of CBI, the results will serve as excellent indicators of the relative success or failure of various parts of the computer-based lessons. However, any information collected with poor tests is worse than useless.

If it is true (as it will be argued later) that few examples of CBI have reliable and valid outcomes measures, then participants` reactions to programs will be the major source of information during operational testing. Programmed questionnaires, euphemistically called "smilometers", are the most frequently used method of collecting participant reaction data during and after CBI. One of the advantages of CBI is that brief questionnaires or individual items can be programmed into the lesson whenever required. However, the evaluator must be alert to the obtrusiveness of his/her questions. Figure 6 is an example of some questions included in a sequence of military-related CBI lessons. In addition to indicators of the participants' global reactions to the CBI, specific information about the perceived difficulty of the lessons, the pace and relevance of the instruction, the reliability of the CBI system, and so forth can be collected. Even more useful information can be gathered with open-ended questions about the best and worst parts of the CBI and specific suggestions for improvement.

The burdensome and time-consuming tasks of processing, analyzing, and reporting this kind of data can be lessened if efforts are made to incorporate data management and analysis programs on the computer. Time invested in setting up such programs before the data is collected will pay off in terms of increasing the timeliness and accuracy of evaluation reports.

Though computer-based questions may be the most frequently used means of collecting participant reaction data, sometimes interviews will yield qualitatively more useful information. Follow-up interviews of CBI users two to six weeks after completion of a sequence of instruction can be particularly revealing. A program component which may have seemed unnoteworthy during CBI may loom large in the minds of participants after they are in another setting for a time. In

addition, in-depth and probing interview questions asked of a few people can often provide better information than the data from a thousand programmed questions.

Another major source of formative evaluation data during operational testing is the observations made by CBI designers and other observers. Figure 7 is an example of a form used by observers to record their observations of students proceeding through CBI designed according to the PLATO system. In addition, CBI can be submitted to systematic review using a package such as MicroSIFT's Evaluator's Guide for Microcomputer- Based Instructional Packages (MicroSIFT, 1982).

To summarize, the second level of evaluation, formative evaluation, is the collection of information for the improvement of CBI. Internal review procedures including expert review and small scale pilot tests are conducted before CBI is actually implemented. Operational testing procedures including computer-based questions, personal interviews, and observations are conducted during or after the actual implementation of CBI. As trite as it sounds, formative evaluation is a continuing process. Though it must be done within the limits of real time and money, formative evaluation is an activity which encourages creativity and innovation for improving CBI, and therefore can be extremely rewarding.

Level 3 - Assessment of Immediate Learner Effectiveness

A CBI project has many different goals and objectives which may be classified as either immediate or long term. Immediate goals/objectives can be accomplished within the timeframe of the project, whereas long term goals/objectives may not be accomplished until some time after the completion of a project. An immediate goal of a CBI project might be to teach decision analysis strategies to medical students, while a long term goal of the same project might be to increase the cost effectiveness of laboratory tests ordered by physicians.

One of the most common methods of evaluating the immediate learner effectiveness of CBI is to include pretests and posttests in the instructional programs themselves. Unfortunately, though these tests are useful as diagnostic and motivational instruments, they cannot be regarded as adequate indicators of the instructional effectiveness of CBI programs. First, such tests are usually too brief to have adequate reliability. Second, the time-lapse between the administration of the pretest and the posttest is often so short that students may choose a correct response after simply eliminating an earlier incorrect response. Thus improvement in test performance may result more from remembering the test than from learning the knowledge and skills necessary to pass the test.

Assessment of immediate learner effectiveness should include comprehensive measurement of a CBI program's terminal skill and knowledge objectives in as realistic a setting as possible. Computer simulations offer the best alternatives to traditional testing procedures. Interactive microcomputer/videodisc simulations can be designed which require students to demonstrate a full range of the learning they have achieved.

A military-related CBI project currently under development at the University of Maryland's Center for Instructional Development and Evaluation provides an excellent example of simulation used to evaluate the immediate learner effectiveness of CBI. The program is intended to teach functional literacy skills to enlisted personnel in military contexts. Rather than present students with pretests and posttests to determine their progress through the lessons in the tired, old lockstep manner of most CBI, students sign up as military "time travelers" who are sent backward and forward into unique times and places. Once there, the soldiers are put into situations which require them to employ their functional literacy skills to solve military-related problems. For example, a soldier student may travel back in time to the "Battle of the Bulge" where he/she has to follow directions and interpret maps to deliver an urgent message to General Patton. Another mission involves a journey to a futuristic city where the soldier has to locate and read information in his/her Soldier's Manual of Common Tasks to "see, communicate, navigate, shoot, and survive" in an unfamiliar and hostile environment.

A student enters an instructional lesson if and only if his/her performance in the simulation indicates a need for that instruction. In other words, the problems posed by the simulation serve as the pretests to determine the instructional needs of the students and as the posttests to assess whether they've learned the required skills. These strategies not only completely individualize the instruction according to the unique functional literacy needs of each student, they also permit the testing and retesting of the functional literacy skills, thus yielding a meaningful evaluation of the instructional effectiveness of the lessons. During a student's first trip or mission, he/she may require a considerable number of the "lessons," but this number should substantially decrease from journey to journey.

Another advantage of using simulations to evaluate the immediate instructional effectiveness of CBI is that the negative effects of traditional testing is avoided. Many students, especially adult learners, have come to expect failure in a testing situation. Computer simulations, which can "test" just as completely as any multiple choice or open-ended questions, permit students to display their knowledge and skills without anxiety. In the same functional literacy CBI project described above, students report that the computer gives them a new lease on their educational lives. This kind of fresh start for students with a long history of failure is invaluable.

Of course, there are situations in which traditional tests are warranted. If reliable and valid tests matching the objectives of a particular CBI program are available, there is no reason why they can not be used for estimating lesson effectiveness. All too often, however, tests do not reflect the actual learning objectives of CBI or they are administered in a manner which invalidates their results.

Whatever types of measures of the immediate instructional effectiveness (simulations, tests, checklists, rating scales) are used, the evaluator should also make some effort to measure the actual implementation of the CBI and relate this information to the outcome data. CBI is usually delivered at a number of different sites under a variety of conditions. The implementation or actual use of a particular CBI program may vary considerably from site to site. The assumption that a student at site A experiences the same instructional program as a student at site B should not go untested. Environmental factors such as hardware reliability, physical setting, power fluctuations, computer system load, and so forth can greatly affect the implementation of CBI. Personal factors such as local support by teachers or administrators can have similar impact on the kind of CBI a student actually experiences. A full discussion of the importance of assessing the actual implementation of CBI (versus the planned implementation) and the procedures for relating this information to outcome data cannot be accomplished in this paper. The interested reader is referred to Reeves (1979).

To summarize, assessment of the immediate learner effectiveness of CBI involves measuring the degree to which the short term learning objectives of CBI have been accomplished. Computer simulations and tests (both of which must meet acceptable standards of reliability and validity) and measures of the actual implementation of CBI are used for this level of evaluation. Obviously, assessing the immediate instructional effectiveness of CBI is a time-consuming and expensive undertaking. If reliable and valid measures of program outcomes and processes are not available, poor measures should not be substituted. It is better to concentrate on other levels of evaluation than to administer unreliable instruments in an effort to evaluate immediate learner effectiveness.

Level 4 - Impact Evaluation

It is beyond the scope and means of most CBI projects to evaluate definitively the ultimate impact of the CBI programs. For example, CBI used to teach functional literacy to soldiers will not only influence the lives and job performance of individual soldiers but will also affect the overall performance of their military units. Furthermore, many CBI projects result in advances in the technology and delivery of CBI which may ultimately impact on other CBI projects. Unfortunately, the current state of the art of evaluation does not permit the

adequate measurement of these kinds of impacts nor their conclusive attribution to a specific project. Still another problem is that the timespan of a CBI project often ends before impacts on the participants or the contexts in which they work can be observed.

Nevertheless, some effort must be made to assess the long term impact of CBI. Two approaches are recommended. First, interviews and anecdotal records can be used to answer questions about the transfer of the knowledge and skills learned through CBI to other environments. These procedures can also be used to collect multiple perspectives on the intended and unintended outcomes of a particular project. Second, a model for relating program inputs and processes to program outcomes can be used in an attempt to explain the costs and effects of a CBI project so that the findings can be used in the development of other CBI projects.

Interview protocols are adequately described in other sources (Benjamin, 1974; Guba & Lincoln, 1981) and anecdotal records are described above. (See Figure 4.) These qualitative data collection procedures can be used to estimate the impact of a CBI project from the multiple perspectives of all the participants in a CBI project. Developers and site managers can use anecdotal records to record revealing stories of the long range influence of CBI projects on users. For example, an anecdote from a military CBI project describes the case of a particularly unmotivated soldier who became so involved in using a computer for instructional purposes that his commanders noted a radical change in work attitude. Interviews conducted six weeks to six months after participation in CBI can be used to estimate the relevance of knowledge, skills, and attitudes acquired through CBI from the viewpoints of users. In a training situation, interviews can be employed to collect supervisors' opinions of the relevance and transfer of learning through CBI. Though this kind of "soft" data is out of favor in some circles, it can be reported in reliable and meaningful formats so as to affect the decision making of program sponsors and other decision makers.

As noted above, a second approach to impact evaluation involves attempts to explain the effects of a CBI project in terms of a project's inputs and processes. This is not something to be undertaken casually, but it can be very rewarding. Evaluation theorists, practitioners, and users have all recently called for comprehensive evaluations which relate program inputs and processes to program outcomes. Very few projects have successfully implemented this approach to evaluation.

Though it is obvious that the outcomes of most instructional programs (CBI or otherwise) can be explained in terms of the abilities and prior achievements students bring to the program and the instructional activities in which they participate during the programs, many if not most evaluations concentrate solely on

measuring program outcomes. A notable exception to outcomes-oriented evaluations is the Instructional Dimensions study conducted by Cooley and Leinhardt (1980). These researchers used a simple model of classroom processes (see figure 8) to evaluate the effectiveness of several reading and math programs in elementary schools. Their approach enabled Cooley and Leinhardt to indicate not only the relative effectiveness of the various programs they examined, but also allowed them to specify the particular components of the programs which explained their relative effectiveness.

Leinhardt (1980) lists the requirements for this type of evaluation as methods of measuring processes as well as inputs and outcomes, causal models of the influence of process on outcomes, and techniques for analyzing the data. She points out that the measurement of instructional processes is often very difficult and expensive. Fortunately, CBI, unlike many other instructional treatments, permits the accurate and efficient collection of instructional process data. CBI, because of the data collection ability of the computer itself, provides a unique opportunity to conduct evaluation research which is not only relevant to explaining the effects of a particular CBI program, but also significant for research on CBI in general. Techniques for managing and analyzing this type of evaluation data are also available. What is missing is a causal model which relates student characteristics and the effective dimensions of CBI to outcomes. The authors are currently constructing a model using the techniques of meta-analysis (Glass, 1976; McGaw & Glass, 1980). When this work is completed, an attempt to validate the model will be made using the data from a large scale military CBI project.

In summary, impact evaluation is the process of assessing the long term effects of CBI. Despite the difficulties described above, some efforts to evaluate the long term impacts of CBI must be made. The sponsors and consumers of CBI should be as informed as possible about the ultimate effects of CBI when they make decisions about continuance, expansion, or selection. Furthermore, resources for CBI will probably always be less than desirable, and the competition for these resources should be based on impact data rather than politics or whimsy.

Conclusion

As noted above, a review of CBI literature reveals that very few CBI projects have employed more than one or two of the four levels of evaluation: documentation, formative evaluation, evaluation of immediate learner effectiveness, and impact evaluation. Partially, this is due to time and budget constraints. In addition, it is suggested that many CBI developers have not considered the full range of evaluation levels applicable in CBI projects. This paper provides CBI project personnel with guidelines and a few sample instruments for the systematic evaluation of the development and implementation of CBI. Figure 9 (adapted from Anderson and Ball, 1978) illustrates the various evaluation methods which can be employed at each level of evaluation.

The developers of CBI have a unique opportunity, if not a responsibility, to thoroughly evaluate their efforts. Unlike many other approaches to instruction, CBI uses a teacher which can be programmed to systematically record every response of the student as well as every instructional strategy of the "teacher." Used systematically and creatively, the instructional monitoring power of computers can yield accurate, timely information which can influence the decision making of project participants in ways impossible in many other education and training contexts.

Resource or Cost Categories / Life Cycle Stages	Category 1	Category 2	Category 3	Category 4...n	Total Costs
Design and Development					$
Production and Delivery					$
Implementation and Operation					$
Total Costs:	$	$	$	$	

Key:

Life Cycle Stages: Dividing project activities into key periods of resource consumption corresponding to the characteristic nature of the resource requirements of each stage:

Design and Development - generally a fixed, nonrecurring investment in planning design and development activities.

Production and Delivery - capital investment in goods and services needed to put plans into action. Generally this means a variable, nonrecurring cost which fluctuates according to scale of implementation and time horizon.

Implementation and Operation - variable, recurring costs whose size varies with scale and cycles of operation.

Cost Center: A cost center refers to that which is being costed - a project, course, organizational unit, etc. The particular cost centers to be employed in a given study depend on the purpose of the analysis and the concerns of the decision-makers.

Resource or Cost Category: While the life cycle and cost center dimensions of the model represent project or program functions, the resource or cost category dimension represents the specific type of resources being consumed. All resources assigned to or consumed by the cost center are organized into categories. The consumption of that resource in its natural units (like hours, # of pieces of hardware) and the dollar value of those units are reported. Typical categories for analyses of instructional technology projects include: personnel salaries and benefits, administrative time, services, hardware, software, and facilities. See Figure VI-4 for further cost category descriptions relevant to this project.

Figure 1. Resource Analysis Model For A Single Cost Center

CBIP ACTIVITY LOG

NAME _____ DATE(S) _____

COST CENTER CODES

1 = EXISTING LESSONS

2 = NEW GENERAL MILITARY LESSONS

3 = NEW MOS SPECIFIC LESSONS

ACTIVITY CODES

DEVELOPMENT CYCLE

1.1 Needs/Task Analysis & Hierarchies
1.2 Courseware Development/Program's
1.3 Software Development
1.4 Videodisc Production-Premaster
1.5 Audio Production-Original/Scripts
1.6 Hardware Development
1.7 Evaluation
1.8 Project Administration

PRODUCTION/DELIVERY CYCLE

2.1 Hardware/Software Purchase
2.2 Media Production/Replication
2.3 Distribution and Training

OPERATION CYCLE

3.1 Hardware Maintenance
3.2 Software Maintenance
3.3 Courseware Maintenance
3.4 Instruction
3.5 Program Administration

DATE	HOURS	COST CENTER	ACTIVITY	EXPLANATION

SITE MANAGER'S INCIDENT REPORT FORMS

INSTRUCTIONS

During the operation of the Computer-Based Instruction Project, you are requested to report any unusual occurrences which may influence this project. Please use the attached forms to record all deviations from normal operating procedures, e.g, equipment breakdowns or major declines in the student population. At the same time, if the problem requires assistance from the CBI office, call the University of Maryland, Heidelberg.

It is important that you record all incidents whether or not you have to call Heidelberg. Please make your explanation of the incident as complete as possible so that the CBI Project personnel will understand the problem. For example, if the system breaks down in the middle of a lesson, please describe exactly what appeared on the screen when the problem occured. Or if the problem involves a decline in student population, describe the reasons, e.g., maneuvers or PCS.

An example of a complete form is below. These forms will be collected during site visits by project personnel. Thank you for your assistance.

SITE MANAGER'S INCIDENT REPORT FORM

DATE: 09 September 1981 TIME: 10:30 AM

INCIDENT: A student was in the middle of lesson 6304 when the computer made a funny noise and "A>" appeared on the screen.

EXPLANATION called Heidelberg. They said to restart the system from the beginning, without doing the SUBMITS Then Check the student's file using Editor.

ACTION TAKEN: Did what they suggested and everything's working again.

Figure 3. Incident Report Form.

Anecdotal Record Form Instructions

As a member of the CBI Project, you will observe incidents or listen to reports of incidents which relate to the development and impact of the project. It is important that this kind of anecdotal information be systematically recorded so that the story of the development and outcomes of this project can be written. Therefore, you are requested to complete one of the attached Anecdotal Record Forms whenever you witness or hear of a significant incident related to the project's progress and accomplishments. An anecdotal record is a verbal account that exhibits these characteristics:

1. Each anecdote should be limited to a single incident.

2. It should contain a factual, non-inferential description of the observed or reported incident. (For example, The soldier said "I've never enjoyed learning anything before." instead of "The soldier expressed satisfaction with the learning situation.")

3. It should include a description of the situation in which the incident occurs so that the meaning of the behavior can be understood.

4. It should be written as soon as possible after witnessing or hearing about the incident so that all the important details can be included.

5. It should include a separate section describing your interpretation or feelings about the anecdote. Your personal evaluation is important because your judgements about the project are valued highly.

Figure 4. Anecdotal Record Form

ANECDOTAL RECORD FORM

DATE_____ PLACE_____

NAME OF OBSERVER_____

Description of the incident:

Interpretation:

Figure 4. (Continued)

Instructions: The following 25 items represent both instructional design and technical design criteria for a well-developed CAI lesson. As you review your lesson, check for the presence or absence of these attributes. A negative response signals need for a possible modification of the lesson.

Yes	No		
( )	( )	1.	Specific skills, knowledge, or abilities required for a learner to interact with the lesson have been specified.
( )	( )	2.	The learner knows what is expected of him/her; that is, objectives have been clearly identified.
( )	( )	3.	A test is included at the beginning of the lesson for diagnostic purposes.
( )	( )	4.	The lesson is structured so that a learner may interact with all or part of the lesson as appropriate for his/her abilities.
( )	( )	5.	The lesson presents new information in a context that directly relates what the learner already knows to the new material.
( )	( )	6.	The lesson is organized so that the learner acquires basic skills before attempting to demonstrate more advanced skills.
( )	( )	7.	A variety of explanations has been presented.
( )	( )	8.	Exercises, problems, or questions are provided for the learner to practice the types of skills, attitudes, or knowledge specified by the objectives.
( )	( )	9.	The lesson is written in such a way as to provide clues to key concepts (use of asterisks, underlining, etc.).
( )	( )	10.	Outlines, summaries, or reviews are provided to help the student organize key ideas.
( )	( )	11.	A variety of different questioning formats (multiple choice, matching, true-false, completion) are utilized.
( )	( )	12.	Questions elicit responses relevant to the content being presented.
( )	( )	13.	Restatement of important concepts is provided to learners to reinforce learning.
( )	( )	14.	When a learner answers incorrectly, feedback is provided to suggest what information appropriate responses should include.
( )	( )	15.	Vocabulary appropriate for the learner has been used.
( )	( )	16.	A post-test is included to determine learner achievement of stated objectives.
( )	( )	17.	The learner knows what to do; any specific procedures for interacting with the computer/terminal have been explained.
( )	( )	18.	Opportunities for frequent interaction between learner and computer have been provided.
( )	( )	19.	The learner is informed about his/her status (score, how many lessons completed, etc.).
( )	( )	20.	The typing skills necessary to make a response have been minimized.
( )	( )	21.	Flexibility in accepting learner responses, especially synonyms, has been developed.
( )	( )	22.	Opportunities for assistance within the lesson (as HELP, RESTART, and REVIEW) are available.
( )	( )	23.	The lesson is not so lengthy as to be tiring.
( )	( )	24.	Any supplementary materials necessary for the use of the lesson are provided for the learners.
( )	( )	25.	The lesson is different than a textbook, lecture, or programmed instruction book.

Figure 5. CAI Readiness Checklist

Strand 1

1. The introductory lesson made it easy to use the computer.

   YES NO NOT SURE

2. The computer is difficult to use.

   YES NO NOT SURE

3. I think I can learn a lot with this computer.

   YES NO NOT SURE

4. I would rather learn this material in a regular class than with this computer.

   YES NO NOT SURE

5. Have you ever learned from a computer before?

   YES NO NOT SURE

Strand 2

1. I like using the computer.

   YES MOST OF THE TIME SOME OF THE TIME NO

2. I am learning a lot with this computer.

   YES NO NOT SURE

3. I would rather learn this material in a regular class than with this computer.

   YES NO NOT SURE

4. The lessons challenge me to do my best work.

   YES MOST OF THE TIME SOME OF THE TIME NO

5. Taking the tests helped me know if I really understood the material.

   YES MOST OF THE TIME SOME OF THE TIME NO

6. I would like to take more lessons on the computer.

   YES NO NOT SURE

Figure 6. Computer-Based Evaluation Questions

7. The computer was available when I wanted to use it.

   YES MOST OF THE TIME SOME OF THE TIME NO

8. These lessons have encouraged me to improve my basic skills.

   YES NO NOT SURE

9. The computer allowed me to work at my own pace.

   YES MOST OF THE TIME SOME OF THE TIME NO

10. I tried to just finish the lessons rather than learn the material.

    YES MOST OF THE TIME SOME OF THE TIME NO

---

11. The computer always seemed to be breaking down.

    USUALLY SOMETIMES NEVER NOT SURE

12. Someone was available to help me when there were computer failures or other problems.

    USUALLY SOMETIMES NEVER NOT SURE

13. For you, most of the work in these lessons was:

    VERY DIFFICULT DIFFICULT ABOUT RIGHT TOO EASY

14. Do you think the Army should continue to develop lessons for the computer?

    YES NO NOT SURE

15. I can use what I learned from these lessons on the job.

    YES NO NOT SURE

Figure 6. (Continued)

For Office Use Only
Reactor Operator Curriculum Module _____
Plato File Name _____

_____
Name

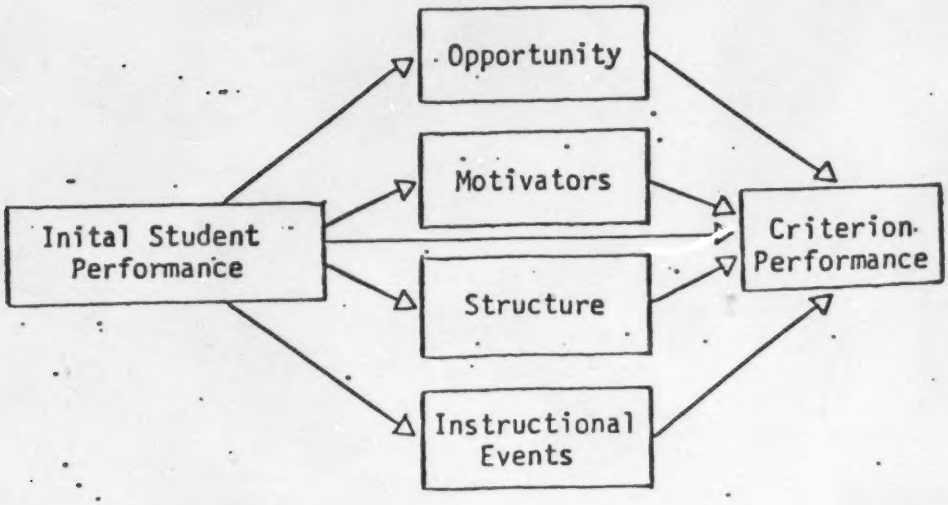
_____
Date

STUDENT USER EVALUATION FORM A

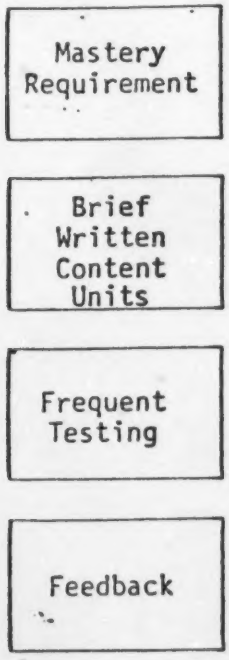
While reviewing the lesson, use this form to write down any problems you may have. Also, indicate any section that is particularly interesting or clearly explains the material. Record the screen no. in the first column. You will find the number in the highlighted area in the lower left hand corner of the screen.

Screen No.	Problem/Comment

Figure 7. Observation Form.



BASIC DIMENSIONS OF
MODEL OF CLASSROOM PROCESSES
(Carroll, 1963; Cooley & Lohnes, 1976)



EFFECTIVE FEATURES OF
PERSONALIZED SYSTEM OF INSTRUCTION
(Johnson & Ruskin, 1977)

Figure 8. Basic dimensions of a model of classroom processes and effective features of PSI.

Evaluation Area	Methods								
	Test Scores	Questionnaire or Interview Data	Logs, Anecdotal Records	Observation	Ratings	Computer-Managed Records	Expert Review	Previous Research Findings	Cost Data and Cost Performance Models
Cost Analysis		✓	✓	✓		✓		✓	✓
Documentation			✓			✓			✓
Formative Evaluation	✓	✓	✓	✓		✓	✓		
Instructional Effectiveness	✓	✓	✓			✓			
Impact	✓	✓	✓		✓	✓	✓		
Cost-Effectiveness	✓	✓	✓	✓	✓	✓	✓	✓	✓

Figure 9. Evaluation Level by Method Matrix.

References

Anderson, S.B. & Ball, S. The profession and practice of program evaluation. San Francisco: Jossey-Bass, Inc., 1978.

Beilby, A. "Determining instructional costs through functional cost analysis." Journal of Instructional Development, 3, (2), 1979.

Benjamin, A. The helping interview (2nd ed.). Boston: Houghton Mifflin Company, 1974.

Cooley, W. W., & Leinhardt, G. "The instructional dimensions study." Educational Evaluation and Policy Analysis, 1980, 2, (1), 7-25.

Felden, J., and Pearson, P. K. The cost of learning with computers. London: Council for Educational Technology, 1978.

Fitzpatrick, M. J., & Howard, J. A. "Field evaluation of a microprocessor based computer for educational applications" (Abstract). Journal of Computer-Based Instruction, 1976, 2, 103-104.

Fletcher, J. D., & Suppes, P. "The Stanford project on computer-assisted instruction for hearing-impaired students." Journal of Computer-Based Instruction, 1976, 3, 1-12.

Glass, G. V. "Primary, secondary, and meta-analysis of research." Educational Researcher, 1976, 5, 3-8.

Guba, E. G., & Lincoln, Y. S. Effective evaluation. San Francisco: Jossey Bass, Inc., 1981.

Leinhardt, G. "Modeling and measuring educational treatment in evaluation." Review of Educational Research, 1980, 50, (3), 393-420.

Lent, R. M. "A model for applying cost-effectiveness analysis to decisions involving the use of instructional technology." Journal of Instructional Development, 2, (5), 1979.

Lent, R. M. "An examination of the methods of cost-effectiveness analysis as applied to instructional technology." (Doctoral dissertation, Syracuse University, 1980). Dissertation Abstracts International, 1980, 41, 2409-A.

McGaw, B., & Glass, G. V. "Choice of the metric for effect size in meta-analysis." American Educational Research Journal, 1980, 17, (3), 325-337.

McPherson-Turner, C. "CAI readiness checklist: Formative author-evaluation of CAI lessons." Journal of Computer- Based Instruction, 1979, 6, (2), 47-49.

MicroSIFT. Evaluator's Guide for Microcomputer-Based Instructional Packages. Eugene, Oregon: ICCE, 1982.

Muston, R., & Wagstaff, R. "Introducing the computer into teacher education" (Abstract). Journal of Computer-Based Instruction, 1976, 2, 108-109.

Nie, N. H., Hull, C. H., Jenkins, J. G., Steinbrenner, K., & Bent, D. H. Statistical package for the social studies (2nd ed.). New York: McGraw-Hill, 1975.

O'Neil, H. F., Jr. (Ed.). Procedures for instructional systems development. New York: Academic Press, 1979.

Rahmlow, H. F., Fratini, R. C., & Ghesquiere, J. R. Plato (From the Instructional Design Library, 30). New Jersey: Educational Technology Publications, 1980.

Reeves, T. C. "Evaluating the implementation of an instructional design across settings." (Doctoral dissertation, Syracuse University, 1979.) Dissertation Abstracts, 1979, 40.5, 2616-A, (University Microfilms No. 7925615).

Rosenshine, B. V. "Primary grades instruction and student achievement gain." Paper presented at the annual meeting of the American Educational Research Association, New York, 1977.

Rubin, H., Geller, J., & Hanks, J. "Computer simulation as a teaching tool in biology - A case study: The Biological Systems Analysis and Simulation Laboratory of the City University of New York." Journal of Computer-Based Instruction, 1977, 3, 91-96.

Su, S. Y. W., & Eman, A. E. "Teaching the principles of software systems on a minicomputer." Journal of Computer- Based Instruction, 1975, 1, 73-79.

Swisser, K. M. "Automated Flanders Interactional Analysis." Journal of Computer-Based Instruction, 1976, 2, 63-66.